**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, C12Q 1/68, A61K 38/17, C07K 16/28, 14/72

A2

(11) International Publication Number: **WO 00/15793**

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/US99/20958

(22) International Filing Date: 17 September 1999 (17.09.99)

(30) Priority Data:
09/156,513 17 September 1998 (17.09.98) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/156,513 (CIP)
Filed on 17 September 1998 (17.09.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, San Francisco, CA 94132 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: HUMAN GPCR PROTEINS

(57) Abstract

The invention provides human GPCR proteins (HGPRP) and polynucleotides which identify and encode HGPRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HGPRP.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine. |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

# HUMAN GPCR PROTEINS

## TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human GPCR proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, neurological, and immune disorders.

## BACKGROUND OF THE INVENTION

The term receptor describes proteins that specifically recognize other molecules. The category is broad and includes proteins with a variety of functions. The bulk of the proteins termed receptors are cell surface proteins which bind extracellular ligands, leading to cellular responses including growth, differentiation, endocytosis, and immune response. Other proteins termed receptors facilitate the specific transport of proteins across the endoplasmic reticulum membrane and localize enzymes to a particular location in the cell.

G protein coupled receptors (GPCR) are a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines; for lipid mediators of inflammation, peptide hormones, and sensory signal mediators. The GPCR becomes activated when the receptor binds its extracellular ligand. Conformational changes in the GPCR, which result from the ligand-receptor interaction, affect the binding affinity of a G protein to the GPCR intracellular domains. This enables GTP to bind with enhanced affinity to the G protein. Activation of the G protein by GTP leads to the interaction of the G protein $\alpha$ subunit with adenylate cyclase or other second messenger molecule generators. This interaction regulates the activity of adenylate cyclase and hence production of a second messenger molecule, cAMP. cAMP regulates phosphorylation and activation of other intracellular proteins. Alternatively, cellular levels of other second messenger molecules, such as cGMP or eicosinoids, may be upregulated or downregulated by the activity of GPCRs. The G protein $\alpha$ subunit is deactivated by hydrolysis of the GTP by GTPase, and the $\beta$, $\gamma$, and $\alpha$ subunits reassociate. The heterotrimeric G protein then dissociates from the adenylate cyclase or other second messenger molecule generator. Activity of GPCR may also be regulated by phosphorylation of the intra- and extracellular domains or loops.

Visual excitation and the phototransmission of light signals is a signaling cascade in which GPCRs play an important role. The process begins in retinal rod cells with the absorption of light by the photoreceptor rhodopsin, a GPCR composed of a 40-kDa protein, opsin, and a chromophore, 11-cis-retinal. The photoisomerization of the retinal chromophore causes a

conformational change in the opsin GPCR and activation of the associated G-protein, transducin. This activation leads to the hydrolysis of cyclic-GMP and the closure of cyclic-GMP regulated, $Ca^{2+}$-specific channels in the plasma membrane of the rod cell. The resultant membrane hyperpolarization generates a nerve signal. Recovery of the dark state of the rod cell involves the activation of guanylate cyclase leading to increased cyclic-GMP levels and the reopening of the $Ca^{2+}$-specific channels (L. Stryer (1991) J. Biol. Chem. 266:10711-10714).

Glutamate receptors form a group of GPCRs that are important in neurotransmission. Glutamate is the major neurotransmitter in the CNS and is believed to have important roles in neuronal plasticity, cognition, memory, learning and some neurological disorders such as epilepsy, stroke, and neurodegeneration (Watson, S. and S. Arkinstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 130-132). These effects of glutamate are mediated by two distinct classes of receptors termed ionotropic and metabotropic. Ionotropic receptors contain an intrinsic cation channel and mediate fast, excitatory actions of glutamate. Metabotropic receptors are modulatory, increasing the membrane excitability of neurons by inhibiting calcium dependent potassium conductances and both inhibiting and potentiating excitatory transmission of ionotropic receptors. Metabotropic receptors are classified into five subtypes based on agonist pharmacology and signal transduction pathways and are widely distributed in brain tissues.

The vasoactive intestinal polypeptide (VIP) family is a group of related polypeptides whose actions are also mediated by GPCRs. Key members of this family are VIP itself, secretin, and growth hormone releasing factor (GRF). VIP has a wide profile of physiological actions including relaxation of smooth muscles, stimulation or inhibition of secretion in various tissues, modulation of various immune cell activities, and various excitatory and inhibitory activities in the CNS. Secretin stimulates secretion of enzymes and ions in the pancreas and intestine and is also present in small amounts in the brain. GRF is an important neuroendocrine agent regulating synthesis and release of growth hormone from the anterior pituitary (Watson, S. and S. Arkinstall supra, pp. 278-283).

The structure of GPCRs is highly-conserved and consists of seven hydrophobic transmembrane (serpentine) regions, cysteine disulfide bridges between the second and third extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with the G-proteins. The consensus pattern of the G-protein

coupled receptors signature (PS00237; SWISSPROT) is characteristic of most proteins belonging to this superfamily (Watson, S. and S. Arkinstall supra, pp. 2-6).

The discovery of new human GPCR proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, neurological, and immune disorders.

## SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human GPCR proteins, referred to collectively as "HGPRP". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-6, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7-12, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:7-12, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7-12, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of

the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex: and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-6, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HGPRP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HGPRP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-6, and fragments thereof.

## BRIEF DESCRIPTION OF THE TABLES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HGPRP.

Table 2 shows features of each polypeptide sequence including potential motifs, homologous sequences, and methods and algorithms used for identification of HGPRP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, conditions, diseases or disorders associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which Incyte clones encoding HGPRP were isolated.

Table 5 shows the programs, their descriptions, references, and threshold parameters used to analyze HGPRP.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that, as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

### DEFINITIONS

"HGPRP" refers to the amino acid sequences of substantially purified HGPRP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to HGPRP, increases or

prolongs the duration of the effect of HGPRP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HGPRP.

An "allelic variant" is an alternative form of the gene encoding HGPRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HGPRP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HGPRP or a polypeptide with at least one functional characteristic of HGPRP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HGPRP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HGPRP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HGPRP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HGPRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HGPRP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HGPRP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to HGPRP, decreases the amount or the duration of the effect of the biological or immunological activity of HGPRP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of HGPRP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind HGPRP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HGPRP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the

complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HGPRP or fragments of HGPRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HGPRP, by northern analysis is indicative of the presence of nucleic acids encoding HGPRP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding HGPRP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide

from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity". A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements

required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., $C_0t$ or $R_0t$ analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HGPRP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HGPRP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:7-12, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:7-12 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:7-12 from related polynucleotide

sequences. A fragment of SEQ ID NO:7-12 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:7-12 and the region of SEQ ID NO:7-12 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HGPRP, or fragments thereof, or HGPRP itself, may comprise a bodily fluid: an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between

polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of HGPRP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant", when used in the context of a polynucleotide sequence, may

encompass a polynucleotide sequence related to HGPRP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

**THE INVENTION**

The invention is based on the discovery of new human GPCR proteins (HGPRP), the polynucleotides encoding HGPRP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, neurological, and immune disorders.

Table 1 lists the Incyte Clones used to derive full length nucleotide sequences encoding HGPRP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NO) of the amino acid and nucleic acid sequences, respectively. Column 3 shows the Clone ID of the Incyte Clone in which nucleic acids encoding each HGPRP were identified, and column 4, the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones, their corresponding cDNA libraries, and shotgun sequences. The clones and shotgun sequences are part of the consensus nucleotide sequence of each HGPRP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, the identity of each protein; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HGPRP. The first column of Table 3 lists the polypeptide sequence identifiers. The second column lists tissue categories which express HGPRP as a fraction of total tissue categories expressing HGPRP. The third column lists the diseases, disorders, or conditions associated with those tissues expressing HGPRP. The fourth column lists

the vectors used to subclone the cDNA library.

The following fragments of the nucleotide sequences encoding HGPRP are useful in hybridization or amplification technologies to identify SEQ ID NO:7-12 and to distinguish between SEQ ID NO:7-12 and related polynucleotide sequences. The useful fragments are the fragment of SEQ ID NO:7 from about nucleotide 235 to about nucleotide 270; the fragment of SEQ ID NO:8 from about nucleotide 218 to about nucleotide 247; the fragment of SEQ ID NO:9 from about nucleotide 271 to about nucleotide 300; the fragment of SEQ ID NO:10 from about nucleotide 273 to about nucleotide 303; the fragment of SEQ ID NO:11 from about nucleotide 542 to about nucleotide 571; and the fragment of SEQ ID NO:12 from about nucleotide 703 to about nucleotide 735.

The invention also encompasses HGPRP variants. A preferred HGPRP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HGPRP amino acid sequence, and which contains at least one functional or structural characteristic of HGPRP.

The invention also encompasses polynucleotides which encode HGPRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:7-12, which encodes HGPRP.

The invention also encompasses a variant of a polynucleotide sequence encoding HGPRP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HGPRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:7-12 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:7-12. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HGPRP.

It will be appreciated by those skilled in the art that, as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HGPRP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HGPRP, and all such variations are to be

considered as being specifically disclosed.

Although nucleotide sequences which encode HGPRP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HGPRP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HGPRP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HGPRP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HGPRP and HGPRP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HGPRP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:7-12 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium

citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq DNA polymerase (PE Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies. Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 system (Hamilton, Reno NV), DNA ENGINE thermal cycler (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (PE Biosystems). Sequencing is then carried out using either ABI PRISM 373 or 377 DNA sequencing systems (PE Biosystems) or the MEGABACE 1000 DNA sequencing system (Amersham Pharmacia Biotech). The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HGPRP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown

sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR software, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HGPRP may be cloned in recombinant DNA molecules that direct expression of HGPRP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the

inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HGPRP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HGPRP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HGPRP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HGPRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of HGPRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HGPRP, the nucleotide sequences encoding HGPRP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HGPRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HGPRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HGPRP and its initiation

codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HGPRP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HGPRP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus,TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HGPRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HGPRP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HGPRP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HGPRP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HGPRP

may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HGPRP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HGPRP. Transcription of sequences encoding HGPRP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HGPRP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HGPRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HGPRP in cell lines is preferred. For example, sequences encoding HGPRP can be transformed into cell lines using expression vectors which may contain viral origins of replication

and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ or *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HGPRP is inserted within a marker gene sequence, transformed cells containing sequences encoding HGPRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HGPRP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HGPRP and that express HGPRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or

protein sequences.

Immunological methods for detecting and measuring the expression of HGPRP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HGPRP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HGPRP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HGPRP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HGPRP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HGPRP may be designed to contain signal sequences which direct secretion of HGPRP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro"

form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HGPRP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HGPRP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HGPRP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HGPRP encoding sequence and the heterologous protein sequence, so that HGPRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HGPRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably $^{35}$S-methionine.

Fragments of HGPRP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (PE Biosystems). Various fragments of HGPRP may be synthesized separately and then combined to produce the full length

molecule.

**THERAPEUTICS**

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HGPRP and GPCR proteins. In addition, the expression of HGPRP is closely associated with cell proliferative and immune disorders, and with neurological tissues. Therefore, HGPRP appears to play a role in cell proliferative, neurological, and immune disorders. In the treatment of disorders associated with increased HGPRP expression or activity, it is desirable to decrease the expression or activity of HGPRP. In the treatment of disorders associated with decreased HGPRP expression or activity, it is desirable to increase the expression or activity of HGPRP.

Therefore, in one embodiment, HGPRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HGPRP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease,

Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder.

In another embodiment, a vector capable of expressing HGPRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HGPRP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HGPRP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HGPRP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HGPRP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HGPRP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HGPRP may be administered to a subject to treat or prevent a disorder associated increased expression or activity of HGPRP. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds HGPRP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HGPRP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HGPRP may be administered to a subject to treat or prevent a disorder associated

increased expression or activity of HGPRP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HGPRP may be produced using methods which are generally known in the art. In particular, purified HGPRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HGPRP. Antibodies to HGPRP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (e.g., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HGPRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HGPRP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HGPRP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HGPRP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-

hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HGPRP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HGPRP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HGPRP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HGPRP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HGPRP. Affinity is expressed as an association constant, $K_a$, which is defined as the molar concentration of HGPRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The $K_a$ determined for a preparation of polyclonal antibodies, which are

heterogeneous in their affinities for multiple HGPRP epitopes, represents the average affinity, or avidity, of the antibodies for HGPRP. The $K_a$ determined for a preparation of monoclonal antibodies, which are monospecific for a particular HGPRP epitope, represents a true measure of affinity. High-affinity antibody preparations with $K_a$ ranging from about $10^9$ to $10^{12}$ L/mole are preferred for use in immunoassays in which the HGPRP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with $K_a$ ranging from about $10^6$ to $10^7$ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HGPRP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HGPRP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HGPRP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HGPRP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HGPRP. Thus, complementary molecules or fragments may be used to modulate HGPRP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HGPRP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HGPRP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HGPRP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HGPRP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA

molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HGPRP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HGPRP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HGPRP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible

modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HGPRP, antibodies to HGPRP, and mimetics, agonists, antagonists, or inhibitors of HGPRP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using

pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be

permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HGPRP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HGPRP or fragments thereof, antibodies of HGPRP, and agonists, antagonists or inhibitors of HGPRP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the $ED_{50}$ (the dose therapeutically effective in 50% of the population) or $LD_{50}$ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the $LD_{50}/ED_{50}$ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the $ED_{50}$ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 $\mu$g to 100,000 $\mu$g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

**DIAGNOSTICS**

In another embodiment, antibodies which specifically bind HGPRP may be used for the diagnosis of disorders characterized by expression of HGPRP, or in assays to monitor patients being treated with HGPRP or agonists, antagonists, or inhibitors of HGPRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HGPRP include methods which utilize the antibody and a label to detect HGPRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HGPRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HGPRP expression. Normal or standard values for HGPRP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HGPRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of HGPRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HGPRP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide

sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HGPRP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HGPRP, and to monitor regulation of HGPRP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HGPRP or closely related molecules may be used to identify nucleic acid sequences which encode HGPRP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HGPRP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HGPRP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:7-12 or from genomic sequences including promoters, enhancers, and introns of the HGPRP gene.

Means for producing specific hybridization probes for DNAs encoding HGPRP include the cloning of polynucleotide sequences encoding HGPRP or HGPRP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as $^{32}P$ or $^{35}S$, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HGPRP may be used for the diagnosis of disorders associated with expression of HGPRP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult

respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder. The polynucleotide sequences encoding HGPRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HGPRP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HGPRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The

nucleotide sequences encoding HGPRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HGPRP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HGPRP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HGPRP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HGPRP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HGPRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HGPRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HGPRP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HGPRP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HGPRP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such

as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HGPRP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HGPRP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HGPRP, or fragments thereof, and washed. Bound HGPRP is then detected by methods well known in the art. Purified HGPRP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HGPRP specifically compete with a test compound for binding HGPRP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HGPRP.

In additional embodiments, the nucleotide sequences which encode HGPRP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred

specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 09/156,513, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (Qiagen, Valencia CA), or an OLIGOTEX mRNA purification kit (Qiagen). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene, or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

**II. Isolation of cDNA Clones**

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a MAGIC or WIZARD MINIPREPS DNA purification system (Promega); an AGTC MINIPREP purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid. QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from Qiagen. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

**III. Sequencing and Analysis**

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (PE Biosystems) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (PE Biosystems) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Amersham Pharmacia Biotech). In another alternative, the cDNAs were amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the probability the greater the

homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, S. San Francisco CA) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GENBANK primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GENBANK databases (described above), SWISSPROT, BLOCKS, PRINTS, PFAM, and PROSITE.

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:7-12. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

**IV. Northern Analysis**

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GENBANK or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40,

although lower scores may identify related molecules.

The results of northern analyses are reported a percentage distribution of libraries in which the transcript encoding HGPRP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease categories included cancer, inflammation/trauma, fetal, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease expression are reported in Table 3.

## V. Extension of HGPRP Encoding Polynucleotides

The full length nucleic acid sequence of SEQ ID NO:7-12 was produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the DNA ENGINE thermal cycler (MJ Research). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing $Mg^{2+}$, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the

sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the nucleotide sequence of SEQ ID NO:7-12 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

**VI. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:7-12 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-$^{32}$P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing $10^7$ counts per minute of the labeled probe is used in a typical membrane-

based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (NYTRAN PLUS, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the blots for several hours, hybridization patterns are compared.

**VII. Microarrays**

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

**VIII. Complementary Polynucleotides**

Sequences complementary to the HGPRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HGPRP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HGPRP. To inhibit transcription, a complementary oligonucleotide is designed from the most

unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HGPRP-encoding transcript.

**IX. Expression of HGPRP**

Expression and purification of HGPRP is achieved using bacterial or virus-based expression systems. For expression of HGPRP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HGPRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HGPRP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HGPRP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HGPRP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HGPRP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified HGPRP obtained by these methods can be used directly in the following activity assay.

**X. Demonstration of HGPRP Activity**

GPCR activity of HGPRP is determined in a ligand-binding assay using candidate ligand

molecules in the presence of $^{125}$I-labeled HGPRP. HGPRP is labeled with $^{125}$I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate ligand molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HGPRP, washed, and any wells with labeled HGPRP complex are assayed. Data obtained using different concentrations of HGPRP are used to calculate values for the number, affinity, and association of HGPRP with the ligand molecules.

**XI. Functional Assays**

HGPRP function is assessed by expressing the sequences encoding HGPRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 $\mu$g of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 $\mu$g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HGPRP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HGPRP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known

by those of skill in the art. Expression of mRNA encoding HGPRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

**XII. Production of HGPRP Specific Antibodies**

HGPRP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HGPRP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

**XIII. Purification of Naturally Occurring HGPRP Using Specific Antibodies**

Naturally occurring or recombinant HGPRP is substantially purified by immunoaffinity chromatography using antibodies specific for HGPRP. An immunoaffinity column is constructed by covalently coupling anti-HGPRP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HGPRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HGPRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HGPRP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HGPRP is collected.

**XIV. Identification of Molecules Which Interact with HGPRP**

HGPRP, or biologically active fragments thereof, are labeled with $^{125}I$ Bolton-Hunter reagent. (See, e.g., Bolton and Hunter, supra.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HGPRP, washed, and any wells with labeled HGPRP complex are assayed. Data obtained using different concentrations of HGPRP are used to calculate values for the number, affinity, and association of HGPRP with the candidate

molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

# Table 1

| Polypeptide SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|---|---|---|---|---|
| 1 | 7 | 1258981 | MENITUT03 | 1258981H1 (MENITUT03), 1442823R1 (THYRNOT03), 1962119T6 (BRSTNOT04), 2059242R6 (OVARNOT03), SATA01180F1, SARB01556F1, SARA01967F1 |
| 2 | 8 | 1459432 | COLNFET02 | 1459432H1 (COLNFET02), 1459432R1 (COLNFET02), 1459432X12 (COLNFET02), 3001554F6 (TLYMNOT06), SAAC00257R1, SAAB00250R1, SAAB00523R1 |
| 3 | 9 | 2214673 | SINTFET03 | 1482004T1 (CORPNOT02), 2214673H1 (SINTFET03), 3073644H1 (BONEUNT01), 3573501F6 (BRONNOT01), 4618526H1 (BRAYDIT01), 4857037H1 (BRSTTUT22), 5025086H1 (OVARNON03) |
| 4 | 10 | 2488822 | KIDNTUT13 | 153210R6 (THP1PLB02), 2488822H1 (KIDNTUT13), 2488822X308B1 (KIDNTUT13), 2488822X310D1 (KIDNTUT13), 3558664T6 (LUNGNOT31) |
| 5 | 11 | 2705201 | PONSAZT01 | 2705201H1 (PONSAZT01), 2705201X325F1 (PONSAZT01), 1262948X325F1 (SYNORAT05), 3141184H1 (SMCCNOT02), 384797R6 (HYPONOB01) |
| 6 | 12 | 3036563 | PENCNOT02 | 3036563H1 (PENCNOT02), 4457161H1 (HEAADIR01), SZAH00352F1, SZAH02656F1, SZAH01730F1, SZAH03622F1, SZAH01163F1, SZAH02669F1, SZAH00249F1 |

# Table 2

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences | Identification | Analytical Methods |
|---|---|---|---|---|---|---|
| 1 | 441 | S85 T164 T274 S306 S344 T81 S118 T407 Y312 Y387 | N191 N405 | M1-A23, I51-V72, C116-A145, I156-L175, M207-P229, G242-T264, E330-K341 | Metabotropic glutamate GPCR | BLOCKS, HMM, MOTIFS, PRINTS, SPSCAN |
| 2 | 353 | S158 T255 S86 T120 S151 S243 S246 T251 T317 S325 | N13 N16 N23 N58 N84 | I42-V66, P78-M99, W109-I149, V159-L180, T209-L232, V254-T278, Y293-R319 | Somatostatin receptor-like GPCR | BLAST, BLOCKS, HMM, MOTIFS, PFAM, PRINTS, PROFILESCAN |
| 3 | 333 | T60 T218 S89 S172 T224 | N8 N110 N300 | Y44-L74, P62-H83, F109-R131, N143-L164, A231-G255, K278-P304 | Rhodopsin-like GPCR | BLAST, BLOCKS, HMM, MOTIFS, PFAM, PRINTS |
| 4 | 396 | S36 S187 T251 S27 T323 S389 | N7 | I46-P70, Y79-I100, L117-F157, R166-S187, S219-F242, L265-L289, S302-K328 | Rhodopsin-like GPCR | BLAST, BLOCKS, HMM, MOTIFS, PFAM, PRINTS, PROFILESCAN |
| 5 | 403 | S360 S368 S47 T318 S337 S5 T33 S123 T398 | N30 N352 | I57-L78, G94-E117, C122-V151, L162-L181, M198-F220, G233-L255 | Metabotropic glutamate GPCR | BLOCKS, HMM, MOTIFS, PRINTS |
| 6 | 807 | T129 S155 S172 S201 S322 S347 S409 S662 S787 S794 S117 T166 T271 T402 T583 T587 T618 S771 | N88 N110 N127 N281 N392 N424 N443 N505 N647 N785 N798 | N425-T452, I475-W499, A549-L572, F636-N647, Q677-G696, H709-W730 | Secretin-like GPCR | BLAST, BLOCKS, HMM, MOTIFS, PRINTS |

# Table 3

| Nucleotide SEQ ID NO: | Tissue Expression (Fraction of Total) | Disease or Condition (Fraction of Total) | Vector |
|---|---|---|---|
| 7 | Reproductive (0.306)<br>ardiovascular (0.177)<br>Gastrointestinal (0.177) | Cell Proliferation (0.839)<br>Inflammation (0.162) | pINCY |
| 8 | Reproductive (0.333)<br>Urologic (0.222)<br>Developmental (0.111) | Cell Proliferation (0.667)<br>Inflammation (0.222) | pINCY |
| 9 | Nervous (0.286)<br>Cardiovascular (0.190)<br>Hematopoietic/Immune (0.190) | Cell Proliferation (0.428)<br>Inflammation (0.333)<br>Neurological (0.143) | pINCY |
| 10 | Developmental (0.333)<br>Cardiovascular (0.167)<br>Hematopoietic/Immune (0.167) | Cell Proliferation (0.833)<br>Inflammation (0.334) | pINCY |
| 11 | Nervous (0.632)<br>Reproductive (0.158)<br>Musculoskeletal (0.105) | Cell Proliferation (0.368)<br>Inflammation (0.316) | pINCY |
| 12 | Cardiovascular (0.244)<br>Reproductive (0.244)<br>Urologic (0.122) | Cell Proliferation (0.732)<br>Inflammation (0.341) | pINCY |

# Table 4

| Nucleotide SEQ ID NO: | Clone ID | Library | Library Comment |
|---|---|---|---|
| 7 | 1258981 | MENITUT03 | Library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer. |
| 8 | 1459432 | COLNFET02 | Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation. |
| 9 | 2214673 | SINTFET03 | Library was constructed using RNA isolated from small intestine tissue removed from a Caucasian female fetus who died at 20 weeks' gestation. |
| 10 | 2488822 | KIDNTUT13 | Library was constructed using RNA isolated from kidney tumor tissue removed from a 51-year-old Caucasian female during a nephroureterectomy. Pathology indicated a grade 3 renal cell carcinoma. Patient history included depressive disorder, hypoglycemia, and uterine endometriosis. Family history included calculus of the kidney, colon cancer, and type II diabetes. |
| 11 | 2705201 | PONSAZT01 | Library was constructed using RNA isolated from diseased pons tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease. |
| 12 | 3036563 | PENCNOT02 | Library was constructed using RNA isolated from penis right corpus cavernosum tissue removed from a male. |

# Table 5

| Program | Description | Reference | Parameter Threshold |
|---|---|---|---|
| ABI FACTURA | A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | |
| ABI/PARACEL FDF | A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA. | Mismatch <50% |
| ABI AutoAssembler | A program that assembles nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | |
| BLAST | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx. | Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402. | *ESTs:* Probability value= 1.0E-8 or less<br>*Full Length sequences:* Probability value= 1.0E-10 or less |
| FASTA | A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch. | Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489. | *ESTs:* fasta E value=1.06E-6<br>*Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less<br>*Full Length sequences:* fastx score=100 or greater |
| BLIMPS | A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions. | Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424. | Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less |
| PFAM | A Hidden Markov Models-based application useful for protein family search. | Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322. | Score=10-50 bits, depending on individual protein families |

# Table 5 cont.

| Program | Description | Reference | Parameter Threshold |
|---|---|---|---|
| ProfileScan | An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite. | Gribskov, M et al (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221. | Score= 4.0 or greater |
| Phred | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability. | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194. | |
| Phrap | A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA. | Score= 120 or greater; Match length= 56 or greater |
| Consed | A graphical tool for viewing and editing Phrap assemblies | Gordon, D. et al. (1998) Genome Res. 8:195-202. | |
| SPScan | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides. | Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439. | Score=5 or greater |
| Motifs | A program that searches amino acid sequences for patterns that matched those defined in Prosite. | Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI. | |

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, and SEQ ID NO:6, and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, and SEQ ID NO:12, and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of HGPRP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of HGPRP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
BANDMAN, Olga
LAL, Preeti
TANG, Y. Tom
CORLEY, Neil C.
GUEGLER, Karl J.
GORGONE, Gina A.
BAUGHN, Mariah R.

<120> HUMAN GPCR PROTEINS

<130> PF-0597 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/156,513
<151> 1998-09-17

<160> 12

<170> PERL Program

<210> 1
<211> 441
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1258981CD1

<400> 1

```
Met Ala Ile His Lys Ala Leu Val Met Cys Leu Gly Leu Pro Leu
  1               5                  10                  15
Phe Leu Phe Pro Gly Ala Trp Ala Gln Gly His Val Pro Pro Gly
             20                  25                  30
Cys Ser Gln Gly Leu Asn Pro Leu Tyr Tyr Asn Leu Cys Asp Arg
             35                  40                  45
Ser Gly Ala Trp Gly Ile Val Leu Glu Ala Val Ala Gly Ala Gly
             50                  55                  60
Ile Val Thr Thr Phe Val Leu Thr Ile Ile Leu Val Ala Ser Leu
             65                  70                  75
Pro Phe Val Gln Asp Thr Lys Lys Arg Ser Leu Leu Gly Thr Gln
             80                  85                  90
Val Phe Phe Leu Leu Gly Thr Leu Gly Leu Phe Cys Leu Val Phe
             95                 100                 105
Ala Cys Val Val Lys Pro Asp Phe Ser Thr Cys Ala Ser Arg Arg
            110                 115                 120
Phe Leu Phe Gly Val Leu Phe Ala Ile Cys Phe Ser Cys Leu Ala
            125                 130                 135
Ala His Val Phe Ala Leu Asn Phe Leu Ala Arg Lys Asn His Gly
            140                 145                 150
Pro Arg Gly Trp Val Ile Phe Thr Val Ala Leu Leu Leu Thr Leu
```

```
                155                 160                 165
Val Glu Val Ile Ile Asn Thr Glu Trp Leu Ile Ile Thr Leu Val
                170                 175                 180
Arg Gly Ser Gly Glu Gly Gly Pro Gln Gly Asn Ser Ser Ala Gly
                185                 190                 195
Trp Ala Val Ala Ser Pro Cys Ala Ile Ala Asn Met Asp Phe Val
                200                 205                 210
Met Ala Leu Ile Tyr Val Met Leu Leu Leu Leu Gly Ala Phe Leu
                215                 220                 225
Gly Ala Trp Pro Ala Leu Cys Gly Arg Tyr Lys Arg Trp Arg Lys
                230                 235                 240
His Gly Val Phe Val Leu Leu Thr Thr Ala Thr Ser Val Ala Ile
                245                 250                 255
Trp Val Val Trp Ile Val Met Tyr Thr Tyr Gly Asn Lys Gln His
                260                 265                 270
Asn Ser Pro Thr Trp Asp Asp Pro Thr Leu Ala Ile Ala Leu Ala
                275                 280                 285
Ala Asn Ala Trp Ala Phe Val Leu Phe Tyr Val Ile Pro Glu Val
                290                 295                 300
Ser Gln Val Thr Lys Ser Ser Pro Glu Gln Ser Tyr Gln Gly Asp
                305                 310                 315
Met Tyr Pro Thr Arg Gly Val Gly Tyr Glu Thr Ile Leu Lys Glu
                320                 325                 330
Gln Lys Gly Gln Ser Met Phe Val Glu Asn Lys Ala Phe Ser Met
                335                 340                 345
Asp Glu Pro Val Ala Ala Lys Arg Pro Val Ser Pro Tyr Ser Gly
                350                 355                 360
Tyr Asn Gly Gln Leu Leu Thr Ser Val Tyr Gln Pro Thr Glu Met
                365                 370                 375
Ala Leu Met His Lys Val Pro Ser Glu Gly Ala Tyr Asp Ile Ile
                380                 385                 390
Leu Pro Arg Ala Thr Ala Asn Ser Gln Val Met Gly Ser Ala Asn
                395                 400                 405
Ser Thr Leu Arg Ala Glu Asp Met Tyr Ser Ala Gln Ser His Gln
                410                 415                 420
Ala Ala Thr Pro Pro Lys Asp Gly Lys Asn Ser Gln Val Phe Arg
                425                 430                 435
Asn Pro Tyr Val Trp Asp
                440

<210> 2
<211> 353
<212> PRT
<213> Homo sapiens


<220>
<221> misc_feature
<223> Incyte ID No: 1459432CD1

<400> 2
Met Asp Leu Glu Ala Ser Leu Leu Pro Thr Gly Pro Asn Ala Ser
  1               5                  10                  15
Asn Thr Ser Asp Gly Pro Asp Asn Leu Thr Ser Ala Gly Ser Pro
                20                  25                  30
Pro Arg Thr Gly Ser Ile Ser Tyr Ile Asn Ile Ile Met Pro Ser
                35                  40                  45
```

```
Val Phe Gly Thr Ile Cys Leu Leu Gly Ile Ile Gly Asn Ser Thr
                50                  55                  60
Val Ile Phe Ala Val Val Lys Lys Ser Lys Leu His Trp Cys Asn
                65                  70                  75
Asn Val Pro Asp Ile Phe Ile Ile Asn Leu Ser Val Val Asp Leu
                80                  85                  90
Leu Phe Leu Leu Gly Met Pro Phe Met Ile His Gln Leu Met Gly
                95                 100                 105
Asn Gly Val Trp His Phe Gly Glu Thr Met Cys Thr Leu Ile Thr
               110                 115                 120
Ala Met Asp Ala Asn Ser Gln Phe Thr Ser Thr Tyr Ile Leu Thr
               125                 130                 135
Ala Met Ala Ile Asp Arg Tyr Leu Ala Thr Val His Pro Ile Ser
               140                 145                 150
Ser Thr Lys Phe Arg Lys Pro Ser Val Ala Thr Leu Val Ile Cys
               155                 160                 165
Leu Leu Trp Ala Leu Ser Phe Ile Ser Ile Thr Pro Val Trp Leu
               170                 175                 180
Tyr Ala Arg Leu Ile Pro Phe Pro Gly Gly Ala Val Gly Cys Gly
               185                 190                 195
Ile Arg Leu Pro Asn Pro Asp Thr Asp Leu Tyr Trp Phe Thr Leu
               200                 205                 210
Tyr Gln Phe Phe Leu Ala Phe Ala Leu Pro Phe Val Val Ile Thr
               215                 220                 225
Ala Ala Tyr Val Arg Ile Leu Gln Arg Met Thr Ser Ser Val Ala
               230                 235                 240
Pro Thr Ser Gln Arg Ser Ile Arg Leu Arg Thr Lys Arg Val Thr
               245                 250                 255
Arg Thr Ala Ile Ala Ile Cys Leu Val Phe Phe Val Cys Trp Ala
               260                 265                 270
Pro Tyr Tyr Val Leu Gln Leu Thr Gln Leu Ser Ile Ser Arg Pro
               275                 280                 285
Thr Pro Thr Phe Val Tyr Leu Tyr Asn Ala Ala Ile Ser Leu Gly
               290                 295                 300
Tyr Ala Asn Ser Cys Leu Asn Pro Phe Val Tyr Ile Val Leu Cys
               305                 310                 315
Glu Thr Phe Arg Lys Arg Leu Val Leu Ser Val Lys Pro Ala Ala
               320                 325                 330
Gln Gly Gln Leu Arg Ala Val Ser Asn Ala Gln Ala Ala Asp Glu
               335                 340                 345
Glu Arg Thr Glu Ser Lys Gly Thr
               350
```

<210> 3
<211> 333
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2214673CD1

```
<400> 3
Met Trp Ser Cys Ser Trp Phe Asn Gly Thr Gly Leu Val Glu Glu
  1               5                  10                  15
Leu Pro Ala Cys Gln Asp Leu Gln Leu Gly Leu Ser Leu Leu Ser
                20                  25                  30
```

```
Leu Leu Gly Leu Val Val Gly Val Pro Val Gly Leu Cys Tyr Asn
                35                  40                  45
Ala Leu Leu Val Leu Ala Asn Leu His Ser Lys Ala Ser Met Thr
                50                  55                  60
Met Pro Asp Val Tyr Phe Val Asn Met Ala Val Ala Gly Leu Val
                65                  70                  75
Leu Ser Ala Leu Ala Pro Val His Leu Leu Gly Pro Pro Ser Ser
                80                  85                  90
Arg Trp Ala Leu Trp Ser Val Gly Gly Glu Val His Val Ala Leu
                95                 100                 105
Gln Ile Pro Phe Asn Val Ser Ser Leu Val Ala Met Tyr Ser Thr
               110                 115                 120
Ala Leu Leu Ser Leu Asp His Tyr Ile Glu Arg Ala Leu Pro Arg
               125                 130                 135
Thr Tyr Met Ala Ser Val Tyr Asn Thr Arg His Val Cys Gly Phe
               140                 145                 150
Val Trp Gly Gly Ala Leu Leu Thr Ser Phe Ser Ser Leu Leu Phe
               155                 160                 165
Tyr Ile Cys Ser His Val Ser Thr Arg Ala Leu Glu Cys Ala Lys
               170                 175                 180
Met Gln Asn Ala Glu Ala Ala Asp Ala Thr Leu Val Phe Ile Gly
               185                 190                 195
Tyr Val Val Pro Ala Leu Ala Thr Leu Tyr Ala Leu Val Leu Leu
               200                 205                 210
Ser Arg Val Arg Arg Glu Asp Thr Pro Leu Asp Arg Asp Thr Gly
               215                 220                 225
Arg Leu Glu Pro Ser Ala His Arg Leu Leu Val Ala Thr Val Cys
               230                 235                 240
Thr Gln Phe Gly Leu Trp Thr Pro His Tyr Leu Ile Leu Leu Gly
               245                 250                 255
His Thr Gly Ile Ile Ser Arg Gly Lys Pro Val Asp Ala His Tyr
               260                 265                 270
Leu Gly Leu Leu His Phe Val Lys Asp Phe Ser Lys Leu Leu Ala
               275                 280                 285
Phe Ser Ser Ser Phe Val Thr Pro Leu Leu Tyr Arg Tyr Met Asn
               290                 295                 300
Gln Ser Phe Pro Ser Lys Leu Gln Arg Leu Met Lys Lys Leu Pro
               305                 310                 315
Cys Gly Asp Arg His Cys Ser Pro Asp His Met Gly Val Gln Gln
               320                 325                 330
Val Leu Ala
```

<210> 4
<211> 396
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2488822CD1

<400> 4

```
Met Phe Arg Pro Leu Val Asn Leu Ser His Ile Tyr Phe Lys Lys
  1               5                  10                  15
Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys Pro
                20                  25                  30
```

```
Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser Ile
                35                  40                  45
Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys Phe
            50                  55                  60
Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser Glu
        65                  70                  75
Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala Asp
    80                  85                  90
Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp Leu
                95                  100                 105
Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met Glu
            110                 115                 120
Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser Thr
        125                 130                 135
Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys Tyr
    140                 145                 150
Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys Cys
155                 160                 165
Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe Ile
                170                 175                 180
Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn Tyr
            185                 190                 195
Tyr Ala Pro Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp Thr
        200                 205                 210
Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu Gly
    215                 220                 225
Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly Ser
230                 235                 240
Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu Ile
                245                 250                 255
Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe Phe
            260                 265                 270
Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe Val
        275                 280                 285
Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr Ile
    290                 295                 300
Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala Leu
305                 310                 315
Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu Met
                320                 325                 330
Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met Asp
            335                 340                 345
Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val Glu
        350                 355                 360
Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro Asp
    365                 370                 375
Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser Thr
380                 385                 390
Arg Leu Asn Ser Tyr Ser
395
```

<210> 5
<211> 403
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2705201CD1

<400> 5

```
Met Phe Val Ala Ser Glu Arg Lys Met Arg Ala His Gln Val Leu
  1               5                  10                  15
Thr Phe Leu Leu Leu Phe Val Ile Thr Ser Val Ala Ser Glu Asn
             20                  25                  30
Ala Ser Thr Ser Arg Gly Cys Gly Leu Asp Leu Leu Pro Gln Tyr
         35                  40                  45
Val Ser Leu Cys Asp Leu Asp Ala Ile Trp Gly Ile Val Val Glu
     50                  55                  60
Ala Val Ala Gly Ala Gly Ala Leu Ile Thr Leu Leu Leu Met Leu
 65                  70                  75
Ile Leu Leu Val Arg Leu Pro Phe Ile Lys Glu Lys Glu Lys Lys
             80                  85                  90
Ser Pro Val Gly Leu His Phe Leu Phe Leu Leu Gly Thr Leu Gly
         95                 100                 105
Leu Phe Gly Leu Thr Phe Ala Phe Ile Ile Gln Glu Asp Glu Thr
    110                 115                 120
Ile Cys Ser Val Arg Arg Phe Leu Trp Gly Val Leu Phe Ala Leu
    125                 130                 135
Cys Phe Ser Cys Leu Leu Ser Gln Ala Trp Arg Val Arg Arg Leu
    140                 145                 150
Val Arg His Gly Thr Gly Pro Ala Gly Trp Gln Leu Val Gly Leu
    155                 160                 165
Ala Leu Cys Leu Met Leu Val Gln Val Ile Ile Ala Val Glu Trp
    170                 175                 180
Leu Val Leu Thr Val Leu Arg Asp Thr Arg Pro Ala Cys Ala Tyr
    185                 190                 195
Glu Pro Met Asp Phe Val Met Ala Leu Ile Tyr Asp Met Val Leu
    200                 205                 210
Leu Val Val Thr Leu Gly Leu Ala Leu Phe Thr Leu Cys Gly Lys
    215                 220                 225
Phe Lys Arg Trp Lys Leu Asn Gly Ala Phe Leu Leu Ile Thr Ala
    230                 235                 240
Phe Leu Ser Val Leu Ile Trp Val Ala Trp Met Thr Met Tyr Leu
    245                 250                 255
Phe Gly Asn Val Lys Leu Gln Gln Gly Asp Ala Trp Asn Asp Pro
    260                 265                 270
Thr Leu Ala Ile Thr Leu Ala Ala Ser Gly Trp Val Phe Val Ile
    275                 280                 285
Phe His Ala Ile Pro Glu Ile His Cys Thr Leu Leu Pro Ala Leu
    290                 295                 300
Gln Glu Asn Thr Pro Asn Tyr Phe Asp Thr Ser Gln Pro Arg Met
    305                 310                 315
Arg Glu Thr Ala Phe Glu Glu Asp Val Gln Leu Pro Arg Ala Tyr
    320                 325                 330
Met Glu Asn Lys Ala Phe Ser Met Asp Glu His Asn Ala Ala Leu
    335                 340                 345
Arg Thr Ala Gly Phe Pro Asn Gly Ser Leu Gly Lys Arg Pro Ser
    350                 355                 360
Gly Ser Leu Gly Lys Arg Pro Ser Ala Pro Phe Arg Ser Asn Val
    365                 370                 375
Tyr Gln Pro Thr Glu Met Ala Val Val Leu Asn Gly Gly Thr Ile
    380                 385                 390
```

Pro Thr Ala Pro Pro Ser His Thr Gly Arg His Leu Trp
395 400

<210> 6
<211> 807
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3036563CD1

<400> 6
Met Gly Thr Tyr His Cys Ile Phe Arg Tyr Lys Asn Ser Tyr Ser
1 5 10 15
Ile Ala Thr Lys Asp Val Ile Val His Pro Leu Pro Leu Lys Leu
20 25 30
Asn Ile Met Val Asp Pro Leu Glu Ala Thr Val Ser Cys Ser Gly
35 40 45
Ser His His Ile Lys Cys Cys Ile Glu Glu Asp Gly Asp Tyr Lys
50 55 60
Val Thr Phe His Met Gly Ser Ser Ser Leu Pro Ala Ala Lys Glu
65 70 75
Val Asn Lys Lys Gln Val Cys Tyr Lys His Asn Phe Asn Ala Ser
80 85 90
Ser Val Ser Trp Cys Ser Lys Thr Val Asp Val Cys Cys His Phe
95 100 105
Thr Asn Ala Ala Asn Asn Ser Val Trp Ser Pro Ser Met Lys Leu
110 115 120
Asn Leu Val Pro Gly Glu Asn Ile Thr Cys Gln Asp Pro Val Ile
125 130 135
Gly Val Gly Glu Pro Gly Lys Val Ile Gln Lys Leu Cys Arg Phe
140 145 150
Ser Asn Val Pro Ser Ser Pro Glu Ser Pro Ile Gly Gly Thr Ile
155 160 165
Thr Tyr Lys Cys Val Gly Ser Gln Trp Glu Glu Lys Arg Asn Asp
170 175 180
Cys Ile Ser Ala Pro Ile Asn Ser Leu Leu Gln Met Ala Lys Ala
185 190 195
Leu Ile Lys Ser Pro Ser Gln Asp Glu Met Leu Pro Thr Tyr Leu
200 205 210
Lys Asp Leu Ser Ile Ser Ile Gly Lys Ala Glu His Glu Ile Ser
215 220 225
Ser Ser Pro Gly Ser Leu Gly Ala Ile Ile Asn Ile Leu Asp Leu
230 235 240
Leu Ser Thr Val Pro Thr Gln Val Asn Ser Glu Met Met Thr His
245 250 255
Val Leu Ser Thr Val Asn Ile Ile Leu Gly Lys Pro Val Leu Asn
260 265 270
Thr Trp Lys Val Leu Gln Gln Gln Trp Thr Asn Gln Ser Ser Gln
275 280 285
Leu Leu His Ser Val Glu Arg Phe Ser Gln Ala Leu Gln Ser Gly
290 295 300
Asp Ser Pro Pro Leu Ser Phe Ser Gln Thr Asn Val Gln Met Ser
305 310 315
Ser Met Val Ile Lys Ser Ser His Pro Glu Thr Tyr Gln Gln Arg
320 325 330

```
Phe Val Phe Pro Tyr Phe Asp Leu Trp Gly Asn Val Val Ile Asp
                335                 340                 345
Lys Ser Tyr Leu Glu Asn Leu Gln Ser Asp Ser Ser Ile Val Thr
                350                 355                 360
Met Ala Phe Pro Thr Leu Gln Ala Ile Leu Ala Gln Asp Ile Gln
                365                 370                 375
Glu Asn Asn Phe Ala Glu Ser Leu Val Met Thr Thr Thr Val Ser
                380                 385                 390
His Asn Thr Thr Met Pro Phe Arg Ile Ser Met Thr Phe Lys Asn
                395                 400                 405
Asn Ser Pro Ser Gly Gly Glu Thr Lys Cys Val Phe Trp Asn Phe
                410                 415                 420
Arg Leu Ala Asn Asn Thr Gly Gly Trp Asp Ser Ser Gly Cys Tyr
                425                 430                 435
Val Glu Glu Gly Asp Gly Asp Asn Val Thr Cys Ile Cys Asp His
                440                 445                 450
Leu Thr Ser Phe Ser Ile Leu Met Ser Pro Asp Ser Pro Asp Pro
                455                 460                 465
Ser Ser Leu Leu Gly Ile Leu Leu Asp Ile Ile Ser Tyr Val Gly
                470                 475                 480
Val Gly Phe Ser Ile Leu Ser Leu Ala Ala Cys Leu Val Val Glu
                485                 490                 495
Ala Val Val Trp Lys Ser Val Thr Lys Asn Arg Thr Ser Tyr Met
                500                 505                 510
Arg His Thr Cys Ile Val Asn Ile Ala Ala Ser Leu Leu Val Ala
                515                 520                 525
Asn Thr Trp Phe Ile Val Val Ala Ala Ile Gln Asp Asn Arg Tyr
                530                 535                 540
Ile Leu Cys Lys Thr Ala Cys Val Ala Ala Thr Phe Phe Ile His
                545                 550                 555
Phe Phe Tyr Leu Ser Val Phe Phe Trp Met Leu Thr Leu Gly Leu
                560                 565                 570
Met Leu Phe Tyr Arg Leu Val Phe Ile Leu His Glu Thr Ser Arg
                575                 580                 585
Ser Thr Gln Lys Ala Ile Ala Phe Cys Leu Gly Tyr Gly Cys Pro
                590                 595                 600
Leu Ala Ile Ser Val Ile Thr Leu Gly Ala Thr Gln Pro Arg Glu
                605                 610                 615
Val Tyr Thr Arg Lys Asn Val Cys Trp Leu Asn Trp Glu Asp Thr
                620                 625                 630
Lys Ala Leu Leu Ala Phe Ala Ile Pro Ala Leu Ile Ile Val Val
                635                 640                 645
Val Asn Ile Thr Ile Thr Ile Val Val Ile Thr Lys Ile Leu Arg
                650                 655                 660
Pro Ser Ile Gly Asp Lys Pro Cys Lys Gln Glu Lys Ser Ser Leu
                665                 670                 675
Phe Gln Ile Ser Lys Ser Ile Gly Val Leu Thr Pro Leu Leu Gly
                680                 685                 690
Leu Thr Trp Gly Phe Gly Leu Thr Thr Val Phe Pro Gly Thr Asn
                695                 700                 705
Leu Val Phe His Ile Ile Phe Ala Ile Leu Asn Val Phe Gln Gly
                710                 715                 720
Leu Phe Ile Leu Leu Phe Gly Cys Leu Trp Asp Leu Lys Val Gln
                725                 730                 735
Glu Ala Leu Leu Asn Lys Phe Ser Leu Ser Arg Trp Ser Ser Gln
                740                 745                 750
His Ser Lys Ser Thr Ser Leu Gly Ser Ser Thr Pro Val Phe Ser
```

```
                755                 760                 765
Met Ser Ser Pro Ile Ser Arg Arg Phe Asn Asn Leu Phe Gly Lys
                770                 775                 780
Thr Gly Thr Tyr Asn Val Ser Thr Pro Glu Ala Thr Ser Ser Ser
                785                 790                 795
Leu Glu Asn Ser Ser Ser Ala Ser Ser Leu Leu Asn
                800                 805
```

<210> 7
<211> 1824
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1258981CB1

<400> 7
cggctcgagc cctcaccagc cggaaagtac gagtcggctc agcctggagg gacccaacca 60
gagcctggcc tgggagccag gatggccatc cacaaagcct tggtgatgtg cctgggactg 120
cctctcttcc tgttcccagg ggcctgggcc cagggccatg tcccacccgg ctgcagccaa 180
ggcctcaacc ccctgtacta caacctgtgt gaccgctctg gggcgtgggg catcgtcctg 240
gaggccgtgg ctggggcggg cattgtcacc acgtttgtgc tcaccatcat cctggtggcc 300
agcctccccc ttgtgcagga caccaagaaa cggagcctgc tggggaccca ggtattcttc 360
cttctgggga ccctgggcct cttctgcctc gtgtttgcct gtgtggtgaa gcccgacttc 420
tccacctgtg cctctcggcg cttcctcttt ggggttctgt tcgccatctg cttctcttgt 480
ctggcggctc acgtctttgc cctcaacttc ctggcccgga agaaccacgg gccccggggc 540
tgggtgatct tcactgtggc tctgctgctg accctggtag aggtcatcat caatacagag 600
tggctgatca tcaccctggt tcggggcagt ggcgagggcg gccctcaggg caacagcagc 660
gcaggctggg ccgtggcctc cccctgtgcc atcgccaaca tggactttgt catggcactc 720
atctacgtca tgctgctgct gctgggtgcc ttcctggggg cctggcccgc cctgtgtggc 780
cgctacaagc gctggcgtaa gcatggggtc tttgtgctcc tcaccacagc cacctccgtt 840
gccatatggg tggtgtggat cgtcatgtat acttacggca acaagcagca caacagtccc 900
acctgggatg accccacgct ggccatcgcc ctcgccgcca atgcctgggc cttcgtcctc 960
ttctacgtca tccccgaggt ctcccaggtg accaagtcca gcccagagca aagctaccag 1020
ggggacatgt accccacccg gggcgtgggc tatgagacca tcctgaaaga gcagaagggt 1080
cagagcatgt tcgtggagaa caaggccttt tccatggatg agccggttgc agctaagagg 1140
ccggtgtcac catacagcgg gtacaatggg cagctgctga ccagtgtgta ccagcccact 1200
gagatggccc tgatgcacaa agttccgtcc gaaggagctt acgacatcat cctcccacgg 1260
gccaccgcca acagccaggt gatgggcagt gccaactcga ccctgcgggc tgaagacatg 1320
tactcggccc agagccacca ggcggccaca ccgccgaaag acggcaagaa ctctcaggtc 1380
tttagaaacc cctacgtgtg ggactgagtc agcggtggcg aggagaggcg gtcggatttg 1440
gggagggccc tgaggacctg gccccgggca agggactctc caggctcctc ctccccctgg 1500
caggcccagc aacatgtgcc ccagatgtgg aagggcctcc ctctctgcca gtgtttgggt 1560
gggtgtcatg ggtgtcccca cccactcctc agtgtttgtg gagtcgagga gccaacccca 1620
gcctcctgcc aggatcacct cggcggtcac actccagcca aatagtgttc tcggggtggt 1680
ggctgggcag cgcctatgtt tctctggaga ttcctgcaac ctcaagagac ttcccaggcg 1740
ctcaggcctg gatcttgctc ctctgtgagg aacaagggtg cctaataaat acatttctgc 1800
tttattaact cttaaaaaaa aaaa 1824

<210> 8
<211> 2152
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1459432CB1

<400> 8
ttatgtctgg tcgactctga attgggcttg gaggcggcac ggctgccagg ctacggaggt 60
agacccccct cccaactgcg gggcttgcgc tccgggacaa ggtggcaggc gctggaggct 120
gccgcagcct gcgtgggtgg aggggagctc agctcggttg tggcagcatg cgaccggcac 180
tggctggatg gacctggaag cctcgctgct gcccactggt cccaatgcca gcaacacctc 240
tgatggcccc gataacctca cttcggcagg atcacctcct cgcacgggga gcatctccta 300
catcaacatc atcatgcctt cggtgttcgg caccatctgc ctcctgggca tcatcgggaa 360
ctccacggtc atcttcgcgg tcgtgaagaa gtccaagctg cactggtgca acaacgtccc 420
cgacatcttc atcatcaacc tctcggtagt agatctcctc tttctcctgg gcatgccctt 480
catgatccac cagctcatgg gcaatggggt gtggcacttt ggggagacca tgtgcaccct 540
catcacggcc atggatgcca atagtcagtt caccagcacc tacatcctga ccgccatggc 600
cattgaccgc tacctggcca ctgtccaccc catctcttcc acgaagttcc ggaagccctc 660
tgtggccacc ctggtgatct gcctcctgtg ggccctctcc ttcatcagca tcacccctgt 720
gtggctgtat gccagactca tccccttccc aggaggtgca gtgggctgcg gcatacgcct 780
gcccaaccca gacactgacc tctactggtt caccctgtac cagtttttcc tggcctttgc 840
cctgcctttt gtggtcatca cagccgcata cgtgaggatc ctgcagcgca tgacgtcctc 900
agtggccccc acctcccagc gcagcatccg gctgcggaca aagagggtga cccgcacagc 960
catcgccatc tgtctggtct tctttgtgtg ctgggcaccc tactatgtgc tacagctgac 1020
ccagttgtcc atcagccgcc cgacccccac ctttgtctac ttatacaatg cggccatcag 1080
cttgggctat gccaacagct gcctcaaccc gtttgtgtac atcgtgctct gtgagacgtt 1140
ccgcaaacgc ttggtcctgt cggtgaagcc tgcagcccag gggcagcttc gcgctgtcag 1200
caacgctcag gcggctgacg aggagaggac agaaagcaaa ggcacctgat acttcccctg 1260
ccaccctgca cacctccaag tcagggcacc acaacacgcc accgggagag atgctgagaa 1320
aaacccaaga ccgctcggga aatgcaggaa ggccgggttg tgaggggttg ttgcaatgaa 1380
ataaatacat tccatggggc tcacacgttg ctggggaggc ctggagtcag gtttggggtt 1440
ttcagatatc agaaatcccc ttgggggagc aggatgagac ctttggatag aacagaagct 1500
gagcaagaga acatgttggt ttggataacc ggttgcacta tatctgtgag ctctcaaatg 1560
tcttcttccc aaggcaagag gtggaagggt actgactggg tttgtttaaa gtcaggcagg 1620
gctggagtga gcagccaggg ccatgttgca caaggcctga gagacgggaa agggcccgat 1680
cgctctttcc cgcctctcac tggtgcgatg gaaggtggcc tttctcccaa gctggtggat 1740
aatgaaaaat aaagcatccc atctctcggc gttccagcat cctgtcaatt tccctttgc 1800
tctagaggat gcatgtttat ttgaggggat gtggcactga gcccacagga gtaaaagccc 1860
agtttgctag gaggtctgct tactgaaaac aaggagacct ggggtgggtg tggttggggg 1920
tcttaaaact aataaaagct ggggtcgggg ggcttttgca gctctggtga cattctctcc 1980
acggggcaca tttgctcagt cactaatcca gcttgagtgt ccgtgtgttc tgcatgtgca 2040
ggggtcattc tagtgcccgg tgtgttggca tcatcttttt gctctagccc ttcctctcca 2100
aaataaaatc aaataaagga aaatctccac ccaaaaaaaa aaaaaaaaaa gg 2152

<210> 9
<211> 1878
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2214673CB1

<400> 9
cgcacagcgc gcaggtcctc accagagctc tggtggccac ctctgtcccg ccatgctgct 60
caccgacagt ggccagggcc cacagcacca agaggcttgg gccacaaagt aaagggtcgc 120
ggagcctcgc cggccgccat gtggagctgc agctggttca acggcacagg gctggtggag 180
gagctgcctg cctgccagga cctgcagctg gggctgtcac tgttgtcgct gctgggcctg 240
gtggtgggcg tgccagtggg cctgtgctac aacgccctgc tggtgctggc caacctacac 300

agcaaggcca gcatgaccat gccggacgtg tactttgtca acatggcagt ggcaggcctg 360
gtgctcagcg ccctggcccc tgtgcacctg ctcggccccc cgagctcccg gtgggcgctg 420
tggagtgtgg gcggcgaagt ccacgtggca ctgcagatcc ccttcaatgt gtcctcactg 480
gtggccatgt actccaccgc cctgctgagc ctcgaccact acatcgagcg tgcactgccg 540
cggacctaca tggccagcgt gtacaacacg cggcacgtgt gcggcttcgt gtggggtggc 600
gcgctgctga ccagcttctc ctcgctgctc ttctacatct gcagccatgt gtccacccgc 660
gcgctagagt gcgccaagat gcagaacgca gaagctgccg acgccacgct ggtgttcatc 720
ggctacgtgg tgccagcact ggccaccctc tacgcgctgg tgctactctc ccgcgtccgc 780
agggaggaca cgcccctgga ccgggacacg ggccggctgg agccctcggc acacaggctg 840
ctggtggcca ccgtgtgcac gcagtttggg ctctggacgc cacactatct gatcctgctg 900
gggcacacgg gcatcatctc gcgagggaag cccgtggacg cacactacct ggggctactg 960
cactttgtga aggatttctc caaactcctg gccttctcca gcagctttgt gacaccactt 1020
ctctaccgct acatgaacca gagcttcccc agcaagctcc aacggctgat gaaaaagctg 1080
ccctgcgggg accggcactg ctccccggac cacatggggg tgcagcaggt gctggcgtag 1140
gcggcccagc cctcctgggg agacgtgact ctggtggacg cagagcactt agttaccctg 1200
gacgctcccc acatccttcc agaaggagac gagctgctgg aagagaagca ggaggggtgt 1260
ttttcttgaa gtttcctttt tcccacaaat gccactcttg ggccaaggct gtggtccccg 1320
tggctggcat ctggcttgag tctccccgag gcctgtgcgt ctcccaaaca cgcagctcaa 1380
ggtccacatc cgcaaaagcc tcctcgcctt cagcctcctc agcattcagt ttgtcaatga 1440
agtgatgaaa gcttagagcc agtatttata ctttgtggtt aaaatacttg attccccctt 1500
gtttgtttta caaaaacaga tgtttcctag aaaaatgaca aatagtaaaa tgaacaaaac 1560
cctacgaaag aatggcaaca gccagggtgg ccgggccctg ccagtgggcg gcgtgtgcta 1620
gcaaggcctg ccgggtgtgc cgcagtcacc acagggttct gagaacattt cacagaagtg 1680
cctgagacgc ggagacatgg ctggtgttaa atggagctat tcaatagcag tgacgcgctc 1740
tcctcagcca ccaaatgtcc ctgacaccct ccccagcccc cacagataac atcagctgag 1800
gtttttttca gtatgaacct gtcctaaatc aattcctcaa agtgtgcaca aaactaaaga 1860
atataaataa acagaagc 1878

<210> 10
<211> 1804
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2488822CB1

<400> 10
taagtgttaa ctaaaagcat tttattaaat tgtccttcac agaaactcaa tttattaaac 60
catgtataat acatgttcct ttgattgatt attaatttga tatttttagc agcctagaag 120
ggattgaaat ttcaaatatc caacaaagga tgtttagacc tcttgtgaat ctctctcaca 180
tatattttaa gaaattccag tactgtgggt atgcaccaca tgttcgcagc tgtaaaccaa 240
acactgatgg aatttcatct ctagagaatc tcttggcaag cattattcag agagtatttg 300
tctgggttgt atctgcagtt acctgctttg gaaacatttt tgtcatttgc atgcgacctt 360
atatcaggtc tgagaacaag ctgtatgcca tgtcaatcat ttctctctgc tgtgccgact 420
gcttaatggg aatatattta ttcgtgatcg gaggctttga cctaaagttt cgtggagaat 480
acaataagca tgcgcagctg tggatggaga gtactcattg tcagcttgta ggatctttgg 540
ccattctgtc cacagaagta tcagttttac tgttaacatt tctgacattg gaaaaataca 600
tctgcattgt ctatcctttt agatgtgtga gacctggaaa atgcagaaca attacagttc 660
tgattctcat ttggattact ggttttatag tggctttcat tccattgagc aataaggaat 720
ttttcaaaaa ctactatgca cccaatggag tatgcttccc tcttcattca gaagatacag 780
aaagtattgg agcccagatt tattcagtgg caatttttct tggtattaat ttggccgcat 840
ttatcatcat agtttttttcc tatggaagca tgttttatag tgttcatcaa agtgccataa 900
cagcaactga aatacggaat caagttaaaa aagagatgat ccttgccaaa cgtttttttct 960
ttatagtatt tactgatgca ttatgctgga tacccatttt tgtagtgaaa tttctttcac 1020
tgcttcaggt agaaatacca ggtaccataa cctcttgggt agtgattttt attctgccca 1080
ttaacagtgc tttgaaccca attctctata ctctgaccac aagaccattt aaagaaatga 1140

ttcatcggtt ttggtataac tacagacaaa gaaaatctat ggacagcaaa ggtcagaaaa 1200
catatgctcc atcattcatc tgggtggaaa tgtggccact gcaggagatg ccacctgagt 1260
taatgaagcc ggacctttтc acataccccт gtgaaatgtc actgatttct caatcaacga 1320
gactcaattc ctattcatga ctgactctga aattcatttc ttcgcagaga atactgtggg 1380
ggtgcttcat gagggattta ctggtatgaa atgaatacca caaaattaat ttataataat 1440
agctaagata aatattttac aaggacatga ggaaaaataa aaatgactaa tgctcttaca 1500
aagggaagta attatatcaa taatgtatat atattagtag acattttgca taagaaatta 1560
agagaaatct acttcagtaa cattcattca tttttctaac atgcatttat tgagtaccca 1620
ctactatgtg catagcattg caatatagtc ctggaagtag acagtgcaga acctttcaat 1680
ctgtagatgg tgtttaatga caaaagacta tacaaagtcc atctgcagtt cctagtttaa 1740
agtagagctt tacctgtcat gtgcatcagc aagaatcata gcgattttaa atagaggtgt 1800
ggac 1804

<210> 11
<211> 1520
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2705201CB1

<400> 11
tgccgaagag tctggagcgt cggcgctgcg gggccgcggg ggtcgaatgt tcgtggcatc 60
agagagaaag atgagagctc accaggtgct caccttcctc ctgctcttcg tgatcacctc 120
ggtggcctct gaaaacgcca gcacatcccg aggctgtggg ctggacctcc tccctcagta 180
cgtgtccctg tgcgacctgg acgccatctg gggcattgtg gtggaggcgg tggccggggc 240
gggcgccctg atcacactgc tcctgatgct catcctcctg gtgcggctgc ccttcatcaa 300
ggagaaggag aagaagagcc ctgtgggcct ccactttctg ttcctcctgg ggaccctggg 360
cctctttggg ctgacgtttg ccttcatcat ccaggaggac gagaccatct gctctgtccg 420
ccgcttcctc tggggcgtcc tctttgcgct ctgcttctcc tgcctgctga gccaggcatg 480
gcgcgtgcgg aggctggtgc ggcatggcac gggccccgcg ggctggcagc tggtgggcct 540
ggcgctgtgc ctgatgctgg tgcaagtcat catcgctgtg gagtggctgg tgctcaccgt 600
gctgcgtgac acaaggccag cctgcgccta cgagcccatg gactttgtga tggccctcat 660
ctacgacatg gtactgcttg tggtcaccct ggggctggcc ctcttcactc tgtgcggcaa 720
gttcaagagg tggaagctga acggggcctt cctcctcatc acagccttcc tctctgtgct 780
catctgggtg gcctggatga ccatgtacct cttcggcaat gtcaagctgc agcaggggga 840
tgcctggaac gaccccacct tggccatcac gctggcggcc agcggctggg tcttcgtcat 900
cttccacgcc atccctgaga tccactgcac ccttctgcca gccctgcagg agaacacgcc 960
caactacttc gacacgtcgc agcccaggat gcgggagacg gccttcgagg aggacgtgca 1020
gctgccgcgg gcctatatgg agaacaaggc cttctccatg gatgaacaca atgcagctct 1080
ccgaacagca ggatttccca acggcagctt gggaaaaaga cccagtggca gcttggggaa 1140
aagacccagc gctccgttta gaagcaacgt gtatcagcca actgagatgg ccgtcgtgct 1200
caacggtggg accatcccaa ctgctccgcc aagtcacaca ggaagacacc tttggtgaaa 1260
gactttaagt tccagagaat cagaatttct cttaccgatt tgcctccctg gctgtgtctt 1320
tcttgaggga gaaatcggta acagttgccg aaccaggccg cctcacagcc aggaaatttg 1380
gaaatcctag ccaaggggat ttcgtgtaaa tgtgaacact gacgaactga aaagctaaca 1440
ccgactgccc gcccctcccc tgccacacac acagacacgt aataccagac caacctcaat 1500
ccccaccttа aaaaaaaaaa 1520

<210> 12
<211> 2919
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3036563CB1

<400> 12
atcttgatgg agcagaatca gtactgacag tcaagacctc gaccagggag tggaatggga 60
acctatcact gcatatttag atataagaat tcatacagta ttgcaaccaa agacgtcatt 120
gttcacccgc tgcctctaaa gctgaacatc atggttgatc ctttggaagc tactgtttca 180
tgcagtggtt cccatcacat caagtgctgc atagaggagg atggagacta caaagttact 240
ttccatatgg gttcctcatc ccttcctgct gcaaaagaag ttaacaaaaa acaagtgtgc 300
tacaaacaca atttcaatgc aagctcagtt tcctggtgtt caaaaactgt tgatgtgtgt 360
tgtcacttta ccaatgctgc taataattca gtctggagcc catctatgaa gctgaatctg 420
gttcctgggg aaaacatcac atgccaggat cccgtaatag gtgtcggaga gccggggaaa 480
gtcatccaga agctatgccg gttctcaaac gttcccagca gccctgagag tcccattggc 540
gggaccatca cttacaaatg tgtaggctcc cagtgggagg agaagagaaa tgactgcatc 600
tctgccccaa taaacagtct gctccagatg gctaaggctt tgatcaagag cccctctcag 660
gatgagatgc tccctacata cctgaaggat ctttctatta gcataggcaa agcggaacat 720
gaaatcagct cttctcctgg gagtctggga gccattatta acatccttga tctgctctca 780
acagttccaa cccaagtaaa ttcagaaatg atgacgcacg tgctctctac ggttaatatc 840
atccttggca agcccgtctt gaacacctgg aaggttttac aacagcaatg gaccaatcag 900
agttcacagc tactacattc agtggaaaga ttttcccaag cattacagtc aggagatagc 960
cctccattgt ccttctccca aactaatgtg cagatgagca gcatggtaat caagtccagc 1020
cacccagaaa cctatcaaca gaggtttgtt ttcccatact ttgacctctg gggcaatgtg 1080
gtcattgaca agagctacct agaaaacttg cagtcggatt cgtctattgt caccatggct 1140
ttcccaactc tccaagccat ccttgctcag gatatccagg aaaataactt tgcagagagc 1200
ttagtgatga caaccactgt cagccacaat acgactatgc cattcaggat ttcaatgact 1260
tttaagaaca atagcccttc aggcggcgaa acgaagtgtg tcttctggaa cttcaggctt 1320
gccaacaaca caggggggtg ggacagcagt gggtgctatg ttgaagaagg tgatggggac 1380
aatgtcacct gtatctgtga ccacctaaca tcattctcca tcctcatgtc ccctgactcc 1440
ccagatccta gttctctcct gggaatactc ctggatatta tttcttatgt tggggtgggc 1500
ttttccatct tgagcttggc agcctgtcta gttgtggaag ctgtggtgtg gaaatcggtg 1560
accaagaatc ggacttctta tatgcgccac acctgcatag tgaatatcgc tgcctccctt 1620
ctggtcgcca acacctggtt cattgtggtc gctgccatcc aggacaatcg ctacatactc 1680
tgcaagacag cctgtgtggc tgccaccttc ttcatccact tcttctacct cagcgtcttc 1740
ttctggatgc tgacactggg cctcatgctg ttctatcgcc tggttttcat tctgcatgaa 1800
acaagcaggt ccactcagaa agccattgcc ttctgtcttg gctatggctg cccacttgcc 1860
atctcggtca tcacgctggg agccacccag ccccgggaag tctatacgag gaagaatgtc 1920
tgttggctca actgggagga caccaaggcc ctgctggctt tcgccatccc agcactgatc 1980
attgtggtgg tgaacataac catcactatt gtggtcatca ccaagatcct gaggccttcc 2040
attggagaca agccatgcaa gcaggagaag agcagcctgt ttcagatcag caagagcatt 2100
ggggtcctca caccactctt gggcctcact tggggttttg gtctcaccac tgtgttccca 2160
gggaccaacc ttgtgttcca tatcatattt gccatcctca atgtcttcca gggattattc 2220
attttactct ttggatgcct ctgggatctg aaggtacagg aagctttgct gaataagttt 2280
tcattgtcga gatggtcttc acagcactca aagtcaacat ccctgggttc atccacacct 2340
gtgttttcta tgagttctcc aatatcaagg agatttaaca atttgtttgg taaaacagga 2400
acgtataatg tttccacccc agaagcaacc agctcatccc tggaaaactc atccagtgct 2460
tcttcgttgc tcaactaaga acaggataat ccaacctacg tgacctcccg gggacagtgg 2520
ctgtgctttt aaaaagagat gcttgcaaag caatggggaa cgtgttctcg gggcaggttt 2580
ccgggagcag atgccaaaaa gacttttca tagagaagag gctttctttt gtaaagacag 2640
aataaaaata attgttatgt ttctgtttgt tccctccccc tccccccttgt gtgataccac 2700
atgtgtatag tatttaagtg aaactcaagc cctcaaggcc caacttctct gtctatattg 2760
taatatagaa tttcgaagag acattttcac tttttacaca ttgggcacaa agataagctt 2820
tgattaaagt agtaagtaaa aggctaccta ggaaatactt cagtgaattc taagaaggaa 2880
ggaaggaagg aaggagggaa agaagggagg aaaccagga 2919

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, C12Q 1/68, A61K 38/17, C07K 16/28, 14/72

A3

(11) International Publication Number: WO 00/15793

(43) International Publication Date: 23 March 2000 (23.03.00)

(21) International Application Number: PCT/US99/20958

(22) International Filing Date: 17 September 1999 (17.09.99)

(30) Priority Data:
09/156,513 17 September 1998 (17.09.98) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/156,513 (CIP)
Filed on 17 September 1998 (17.09.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, San Francisco, CA 94132 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*

(88) Date of publication of the international search report:
28 September 2000 (28.09.00)

(54) Title: HUMAN GPCR PROTEINS

(57) Abstract

The invention provides human GPCR proteins (HGPRP) and polynucleotides which identify and encode HGPRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HGPRP.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

# INTERNATIONAL SEARCH REPORT

Internatic \pplication No
PCT/US 99/20958

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C12Q1/68 A61K38/17 C07K16/28
C07K14/72

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | "NCI-CGAP" DATABASE EMBL ID AA877534, ACCESSION NUMBER AA877534, 30 March 1998 (1998-03-30), XP002110947 98,4% identity with seq ID no.7 in 625 bp overlap nt.1184-1810 reverse orientation abstract --- -/-- | 5,9-12 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 22 February 2000 | 28. 06. 00 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | LE CORNEC N.D.R. |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Internatic Application No
PCT/US 99/20958

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | M.D. ADAMS ET AL: "EST114556 Uterus tumor I Homo sapiens cDNA 5' end" EMBL DATABASE ENTRY HSZZ08200, ACCESSION NUMBER AA303063, 18 April 1997 (1997-04-18), XP002131040 99,% identity in 284 bp overlap : nt616-nt899from seq ID no. 7. abstract & M.D. ADAMS ET AL: "Rapid cDNA sequencing (Expressed sequence tags) from a directionally cloned human infant brain cDNA library" NATURE GENETICS, vol. 4, 1993, pages 373-380, XP000600257 --- | 9-13 |
| X | L. HILLIER ET AL: "EST. Homo sapiens cDNA clone 663473 5'" EMBL DATABASE ENTRY HS1143823, ACCESSION NUMBER AA224173, 22 February 1997 (1997-02-22), XP002131041 98,1% identity in 413 bp overlap: nt221-nt628 from seq ID no.7 abstract --- | 9-13 |
| A | WO 94 29449 A (SALK INST BIOTECH IND ;DAGGETT LORRIE (US); ELLIS STEVEN B (US); L) 22 December 1994 (1994-12-22) the whole document --- | 1-20 |
| A | S. WATSON AND S. ARKINSTAL: THE G-PROTEIN LINKED RECEPTOR FACTSBOOK,1994, XP002131042 GB, London, Academic press cited in the application page 2 -page 6 page 252 -page 260 page 130 -page 141 --- | 1-20 |
| A | OLIVEIRA L ET AL: "A COMMON MOTIF IN G-PROTEIN-COUPLED SEVEN TRANSMEMBRANE HELIX RECEPTORS" JOURNAL OF COMPUTER-AIDED MOLECULAR DESIGN,XX,ESCOM SCIENCE PUBLISHERS BV, vol. 7, no. 6, 1 December 1993 (1993-12-01), pages 649-658, XP002050853 ISSN: 0920-654X --- | |
| A | WO 96 29404 A (SIBIA NEUROSCIENCES INC ;DAGGETT LORRIE P (US); LU CHIN CHUN (US)) 26 September 1996 (1996-09-26) the whole document --- | 1-20 |

-/--

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Internatic Application No
PCT/US 99/20958

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | WO 99 06550 A (LACROIX BRUNO ;DUCLERT AYMERIC (FR); GENSET (FR); DUMAS MILNE EDWA) 11 February 1999 (1999-02-11)<br>Seq ID no. 220 is 99% identical to seq ID no.7: nt 842-1047.<br>seq ID no. 498 is 100% identical to seq ID no.1: aa 262-320.<br>claims<br>--- | 1-20 |
| E | WO 99 53054 A (SMITHKLINE BEECHAM PLC) 21 October 1999 (1999-10-21)<br>sequences ID 1 and 2<br>claims<br>--- | 1-20 |
| E | WO 99 60121 A (SAVITSKY KINNERET ;COMPUGEN LTD (IL); MINTZ LIAT (IL); TOPORIK AMI) 25 November 1999 (1999-11-25)<br>sequences ID no. 1 and 2 pages 44-49<br>abstract; claims<br>----- | 1-20 |

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/20958

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claim 19 is directed to a method of treatment of the human/animal body (rule 39.1 IV PCT.), the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4 [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-20, all partially.

**Remark on Protest**

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 17-18,20 all partially as far as they do not concern the antibody of the polypeptides

Claims 17,18 and 20 refer to an agonist, an antagonist of the polypeptides and a method of treatment or prevention using the antagonist without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported.
No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-20 all partially

A polypeptide comprising the amino acid sequence ID no. 1 encoded by the polynucleotide sequence ID no. 7. Fragments thereof. Method of production by genetic engineering. Host cell. Antibody. Therapeutic uses.

2. Claims: 1-20 all partially

A polypeptide comprising the amino acid sequence ID no. 2 encoded by the polynucleotide sequence ID no. 8. Fragments thereof. Method of production by genetic engineering. Host cell. Antibody. Therapeutic uses.

3. Claims: 1-20 all partially

A polypeptide comprising the amino acid sequence ID no. 3 encoded by the polynucleotide sequence ID no. 9. Fragments thereof. Method of production by genetic engineering. Host cell. Antibody. Therapeutic uses.

4. Claims: 1-20 all partially

A polypeptide comprising the amino acid sequence ID no. 4 encoded by the polynucleotide sequence ID no. 10. Fragments thereof. Method of production by genetic engineering. Host cell. Antibody. Therapeutic uses.

5. Claims: 1-20 all partially

A polypeptide comprising the amino acid sequence ID no. 6 encoded by the polynucleotide sequence ID no. 12. Fragments thereof. Method of production by genetic engineering. Host cell. Antibody. Therapeutic uses.

XCID: <WO___0015793A3_I_>

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Internatic Application No
PCT/US 99/20958

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9429449 A | 22-12-1994 | US 5521297 A | 28-05-1996 |
| | | AU 685471 B | 22-01-1998 |
| | | AU 7098994 A | 03-01-1995 |
| | | CA 2161811 A | 22-12-1994 |
| | | EP 0701611 A | 20-03-1996 |
| | | GB 2286398 A,B | 16-08-1995 |
| | | JP 8511168 T | 26-11-1996 |
| | | US 6001581 A | 14-12-1999 |
| | | US 5912122 A | 15-06-1999 |
| | | US 5807689 A | 15-09-1998 |
| WO 9629404 A | 26-09-1996 | US 5912122 A | 15-06-1999 |
| | | AU 5314696 A | 08-10-1996 |
| | | CA 2215730 A | 26-09-1996 |
| | | EP 0815227 A | 07-01-1998 |
| | | JP 11503605 T | 30-03-1999 |
| WO 9906550 A | 11-02-1999 | AU 8555198 A | 22-02-1999 |
| | | EP 1000148 A | 17-05-2000 |
| WO 9953054 A | 21-10-1999 | NONE | |
| WO 9960121 A | 25-11-1999 | AU 3952699 A | 06-12-1999 |